Collinear Triple Hypergraphs and the Finite Plane Kakeya Problem

Joshua Cooper

February 2, 2008

Abstract

We show that the problem of counting collinear points in a permutation (previously considered by the author and J. Solymosi in [1]) and the well-known finite plane Kakeya problem are intimately connected. Via counting arguments and by studying the hypergraph of collinear triples we show a new lower bound $(5q/14 + O(1))$ for the number of collinear triples of a permutation of $\mathbb{F}_q$ and a new lower bound $(q(q+1)/2 + 5q/14 + O(1))$ on the size of the smallest Besicovitch set in $\mathbb{F}_q^2$. Several interesting questions about the structure of the collinear triple hypergraph are presented.

1   Introduction

The Kakeya problem, solved in 1928 by Besicovitch, asks for the smallest subset of $\mathbb{R}^2$ so that a unit-length interval may be continuously rotated 360° within it. The striking answer is that there are sets of arbitrarily small measure that permit such a "full turn." If one does not require that the interval be continuously rotated, but simply that the set contain a unit interval in every direction, there are even sets of measure zero that do the job. This statement also holds in higher dimensions. Perhaps unsurprisingly, such sets must still be "large," i.e., have Hausdorff (or Minkowski) dimension equal to that of the ambient space. To show this is the case in dimensions higher than two is a very important open problem, with deep connections to additive number theory and harmonic analysis.

In his influential manuscript [4], Wolff describes the "finite field Kakeya problem" and makes a convincing argument that the problems that arise are very similar to those of the analytic setting. The finite field Kakeya problem can be stated as follows. A *Besicovitch set* in $\mathbb{F}_q^d$ is a subset that contains a line in every direction, i.e., a translate of every line through the origin. What is the smallest Besicovitch set? The natural conjecture is that it has cardinality cq^d, but this is not known in dimensions greater than 2. (Indeed, the best result for dimension 3 is only $q^{5/2+10^{-10}}$ [3].)

It is easy to see that, for the finite plane, every Besicovitch set has cardinality at least $q^2/2$. In fact, this result appears already in [4]. On the other hand, there is a natural construction that gives an upper bound of $q(q+1)/2 + (q-1)/2$ for q odd. Hence, we have the following conjecture.

Conjecture 1. *For q an odd prime power, the smallest Besicovitch set in $\mathbb{F}_q^2$ has cardinality $q(q+1)/2 + (q-1)/2$.*

(The situation when q is a power of two is considerably simpler.) Recently, Faber [2] provided a very elegant argument that, in fact, the smallest finite plane Besicovitch set has cardinality at least $q(q+1)/2 + q/3$, getting quite close to the upper bound.

In the current paper, we show that (1) Faber's argument dualizes to improve the best known bound for the number of collinear triples of a permutation (see [1] and the next section); and (2) with some additional work, it is possible to prove the following theorem:

Theorem 1. *The smallest Besicovitch set in $\mathbb{F}_q^2$ has cardinality at least*

$$\frac{q(q+1)}{2} + \frac{5q}{14} - \frac{1}{14}.$$

Note that $5/14 = 1/3 + 1/42$, so this is indeed an improvement. In fact, this lower bound gives the conjectural upper bound for $q = 3, 5, 7, 9$. We include below several very interesting problems that arise from our analysis.

2 Preliminaries

Let q be an odd prime power. Write $\mathbb{F}_q$ for the finite field of order q and $\mathbb{F}_q^*$ for its multiplicative group, i.e., $\mathbb{F}_q \setminus \{0\}$. Let $f : \mathbb{F}_q \to \mathbb{F}_q$ be some function; then we define the "graph of f", denoted Γ_f, to be the set $\{(j, f(j)) : j \in \mathbb{F}_q\} \subset \mathbb{F}_q^2$.

We will write $m_{ij} = m_{ij}(f)$ for the slope of the line connecting the points in Γ_f corresponding to the ordinates i and j, i.e., $m_{ij} = (f(j) - f(i))/(j - i)$. We refer to the slope of a collinear set of points in the obvious sense, and call two collinear sets with the same slope "parallel." In general, a slope is an element of $PG(1, q)$, and therefore can take on the value ∞ if the denominator of this expression is 0, although this is clearly never the case for a graph Γ_f. Also, the slope of points in Γ_σ cannot be 0 for σ a permutation. Finally, we write $\ell(\mathbf{x}, m)$ for the line in $\mathbb{F}_q^2$ through $\mathbf{x}$ with slope m.

Suppose that P is a minimal Besicovitch set in $\mathbb{F}_q^2$. Clearly, we may assume that P is of the form

$$P = \bigcup_{m \in PG(1,q)} \ell(\mathbf{x}_m, m). \tag{1}$$

Write $\mu(\mathbf{y}) = |\{m \in PG(1, q) : \mathbf{y} \in \ell(\mathbf{x}_m, m)\}|$, i.e., the multiplicity of $\mathbf{y}$ in the representation of P by (1). We have the following "incidence formula" for the size of P from [2]:

Theorem 2 (Incidence Formula, [2]).

$$|P| = \frac{q(q + 1)}{2} + \sum_{\mathbf{y} \in P} \binom{\mu(\mathbf{y}) - 1}{2}.$$

Note that $\binom{\mu(\mathbf{y})}{3} \geq \binom{\mu(\mathbf{y}) - 1}{2} \geq \mu(\mathbf{y}) - 2 \geq 1$ for $\mu(\mathbf{y}) \geq 3$. The following result (actually, an equivalent version) appears in [2]. For a Besicovitch set P of the form (1), define $\rho(m)$ to be the sum of $\mu(\mathbf{y}) - 2$ over all $\mathbf{y} \in \ell(\mathbf{x}_m, m)$ with $\mu(\mathbf{y}) \geq 3$.

Proposition 3. *Given a Besicovitch set* $P = \bigcup_{m \in PG(1,q)} \ell(\mathbf{x}_m, m)$, *at most one slope* $m \in PG(1, q)$ *has* $\rho(m) = 0$.

It is not hard to conclude from this Proposition and Theorem 2 that $|P| \geq q(q + 1)/2 + q/3$. We can go further with this result, however. Indeed, suppose that $\min_{m \in PG(1,q)} \rho(m) = R$. If $R \geq 2$, then

$$|P| = \frac{q(q + 1)}{2} + \sum_{\mathbf{y} \in P} \binom{\mu(\mathbf{y}) - 1}{2}$$

$$= \frac{q(q + 1)}{2} + \sum_{\substack{m \in PG(1,q)}} \sum_{\substack{\mathbf{y} \in \ell(\mathbf{x}_m, m) \\ \mu(\mathbf{y}) \geq 3}} \frac{1}{\mu(\mathbf{y})} \binom{\mu(\mathbf{y}) - 1}{2}$$

3

$$= \frac{q(q+1)}{2} + \sum_{\substack{m \in PG(1,q)}} \sum_{\substack{\mathbf{y} \in \ell(\mathbf{x}_m, m) \\ \mu(\mathbf{y}) \geq 3}} \left(\frac{1}{2} - \frac{1}{2\mu(\mathbf{y})} \right) \cdot (\mu(\mathbf{y}) - 2)$$

$$\geq \frac{q(q+1)}{2} + \sum_{\substack{m \in PG(1,q)}} \sum_{\substack{\mathbf{y} \in \ell(\mathbf{x}_m, m) \\ \mu(\mathbf{y}) \geq 3}} \frac{\mu(\mathbf{y}) - 2}{3}$$

$$= \frac{q(q+1)}{2} + \sum_{\substack{m \in PG(1,q)}} \frac{\rho(m)}{3}$$

$$\geq \frac{q(q+1)}{2} + \frac{2(q+1)}{3}.$$

Hence, Theorem 1 follows unless $R = 0$ or $R = 1$.

Now, we may apply a nonsingular affine transformation of $\mathbb{F}_q^2$ to make the line $\ell(\mathbf{x}_m, m)$ with $\rho(m) = R$ become the vertical line through the origin while retaining the property of P being a minimal Besicovitch set. Dualizing the *nonvertical* lines of P under the map $\{(x, y) : y = mx + b\} \mapsto (m, b)$ gives a set $S \subset \mathbb{F}_q^2$ which is the graph Γ_σ of some function σ. If $R = 0$, then σ is a permutation; if $R = 1$, then σ is a "semipermutation", that is, a function whose range has cardinality $q - 1$. If $R = 0$, points $\mathbf{y}$ with $\mu(\mathbf{y}) \geq 3$ correspond to collinear $\mu(\mathbf{y})$-tuples in Γ_σ; if $R = 1$, then all $\mathbf{y}$ with $\mu(\mathbf{y}) \geq 3$ except the single point of multiplicity ≥ 3 on $\ell((0,0), \infty)$ correspond to collinear $\mu(\mathbf{y})$-tuples in Γ_σ.

This leads us to the following definition. Given $S \subset \mathbb{F}_q^2 = \mathbb{F}_q \times \mathbb{F}_q$, define the 3-uniform hypergraph $\mathcal{C}(S)$ on the vertex set S to be the set of all collinear triples $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\} \subset S$. Let $\mathcal{L}(S)$ for the hypergraph on the vertex set S whose edges are the maximal collinear subsets of S. For a hypergraph H, write $\|H\|$ for the quantity $\sum_{e \in E(H)} (|e| - 1)(|e| - 2)/2$; we write $\|S\| = \|\mathcal{L}(S)\|$ for convenience. Note that the number of edges in $\mathcal{C}(S)$, which we write simply $|\mathcal{C}(S)|$, is at least $\|S\|$. It is clear by the Incidence Formula that, when $R = 0$ and σ is the permutation arising from dualizing P,

$$|P| = \frac{q(q+1)}{2} + \|\Gamma_\sigma\|.$$

When $R = 1$ and σ is the semipermutation arising from dualizing P,

$$|P| = 1 + \frac{q(q+1)}{2} + \|\Gamma_\sigma\|.$$

4

Our problem therefore reduces to that of determining the minimum value of $\|\Gamma_\sigma\|$ when σ is a permutation or a semipermutation. In both cases, we may assume that every point of Γ_σ appears in at least one edge of $\mathcal{C}(\Gamma_\sigma)$ by the minimality of R and Proposition 3. In fact, we reprove this statement below for permutations, since the argument is so short and elegant.

The following conjecture appears in [1], along with a proof of the lower bound $(q-1)/4$.

Conjecture 2. *For any permutation* $\sigma : \mathbb{F}_q \to \mathbb{F}_q$, *the number of collinear triples in* Γ_σ *is at least* $(q-1)/2$.

If this conjecture is true, it is best possible, as evidenced by the graph of the function $s \mapsto s^{-1}$ for $s \neq 0$ and $0 \mapsto 0$. Unfortunately, we still cannot show the full conjecture. However, we do show below that every permutation σ satisfies $\|\Gamma_\sigma\| \geq (5q-1)/14$, from which it follows that the number of collinear triples in Γ_σ is at least this quantity. Furthermore, we show that every semipermutation σ so that $\mathcal{C}(\Gamma_\sigma)$ has no isolated points satisfies $\|\Gamma_\sigma\| \geq 5(q-1)/14$. In the former case, then, $|P| \geq q(q+1)/2 + (5q-1)/14$, and in the latter,

$$|P| \geq \frac{q(q+1)}{2} + \frac{5q-5}{14} + 1 = \frac{q(q+1)}{2} + \frac{5q}{14} + \frac{9}{14}.$$

Theorem 1 follows.

3 Permutations

Throughout this section we assume that $\sigma : \mathbb{F}_q \to \mathbb{F}_q$ is a permutation. The next proposition states that no point of Γ_σ is isolated in $\mathcal{C}(\Gamma_\sigma)$, and the argument is essentially the dual of that appearing in [2].

Proposition 4. *For every* $j \in \mathbb{F}_q$, *there exists an* $E \in \mathcal{C}(\Gamma_\sigma)$ *so that* $(j, \sigma(j)) \in E$.

Proof. Clearly we may assume that $(j, \sigma(j)) = (0,0)$. Suppose $(j, \sigma(j))$ were an isolated point in Γ_σ, i.e., there exists no $E \in \mathcal{C}(\Gamma_\sigma)$ so that $(j, \sigma(j)) \in E$. Consider the slopes m_{0i} for $i \neq 0$, and, in particular, the product

$$M = \prod_{i \in \mathbb{F}_q^*} m_{0i}.$$

Since $(0,0)$ does not participate in any collinear triple, the line $\ell((0,0), m)$ contains at most two points for each $m \in \mathbb{F}_q^*$ (including the origin). However, since there are $q-1$ such lines and $q-1$ points $(j, \sigma(j))$ for $j \neq 0$, each such line contains exactly two points of Γ_σ, by the pigeonhole principle. This means that M is the product of all elements of $\mathbb{F}_q^*$, which it is easy to see is -1. On the other hand,

$$M = \prod_{i \in \mathbb{F}_q^*} \frac{\sigma(i)}{i} = \frac{\prod_{i \in \mathbb{F}_q^*} i}{\prod_{i \in \mathbb{F}_q^*} i} = 1,$$

a contradiction. $\qquad\square$

In fact, Proposition 4 can be pushed considerably further.

Proposition 5. *Suppose that* $\mathbf{x} = (j, \sigma(j))$ *participates in no collinear quadruples. Denote by* $\mathcal{M}_i$, $i = 1, 2, 3$, *the set of slopes* $m \in \mathbb{F}_q^*$ *so that* $\ell(\mathbf{x}; m)$ *contains exactly* i *points of* Γ_σ. *Then* $|\mathcal{M}_1| = |\mathcal{M}_3|$, *and*

$$\prod_{m \in \mathcal{M}_1} m = - \prod_{m \in \mathcal{M}_3} m.$$

Proof. First of all, $|\mathcal{M}_2| + 2|\mathcal{M}_3| = q - 1$, since there are $q - 1$ points in $\Gamma_\sigma \setminus \{\mathbf{x}\}$, and $|\mathcal{M}_1| + |\mathcal{M}_2| + |\mathcal{M}_3| = q - 1$, since there are $q - 1$ slopes of lines through $(j, \sigma(j))$ which are not vertical or horizontal. Subtracting the two equations, we see that $|\mathcal{M}_1| = |\mathcal{M}_3|$.

We may assume without loss of generality that $\mathbf{x} = (0, 0)$. Consider again the product

$$M = \prod_{j \in \mathbb{F}_q^*} \frac{\sigma(j)}{j}.$$

On the one hand, the numerator and denominator of this expression both enumerate all elements of $\mathbb{F}_q^*$, so that $M = 1$. On the other hand,

$$M = \left(\prod_{m \in \mathcal{M}_2} m \right) \left(\prod_{m \in \mathcal{M}_3} m^2 \right)$$

$$= \left(\prod_{m \in \mathbb{F}_q^*} m \right) \left(\prod_{m \in \mathcal{M}_1} m \right)^{-1} \left(\prod_{m \in \mathcal{M}_3} m \right)$$

$$= - \left(\prod_{m \in \mathcal{M}_1} m \right)^{-1} \left(\prod_{m \in \mathcal{M}_3} m \right).$$

Hence,

$$\prod_{m \in \mathcal{M}_1} m = - \prod_{m \in \mathcal{M}_3} m.$$

$\square$

Call a point $p \in \Gamma_\sigma$ which belongs to exactly one collinear triple "lonely," i.e.,

$$|\{E \in \mathcal{C}(\Gamma_\sigma) : p \in E\}| = 1.$$

Proposition 6. *If* $\mathbf{x} = (j, \sigma)$ *is lonely, then*

1. *There is one slope* $m_j \in \mathbb{F}_q^*$ *so that* $\ell(\mathbf{x}, m_j)$ *contains three points of* Γ_σ, *one slope* $m_j^\perp$ *so that* $\ell(\mathbf{x}, m_j^\perp)$ *contains one point of* Γ_σ, *and all other slopes* $m \neq m_j, m_j^\perp$ *are such that* $\ell(\mathbf{x}, m)$ *contains two points.*

2. $m_j^\perp = -m_j$.

Proof. Since $\mathbf{x}$ participates in exactly one collinear triple, it is easy to see that, letting m_j be the slope of that triple, $|\ell(\mathbf{x}, m_j) \cap \Gamma_\sigma| = 3$ and $|\ell(\mathbf{x}, m) \cap \Gamma_\sigma| \leq 2$ for all other $m \in \mathbb{F}_q^* \setminus \{m_j\}$. There are $q - 2$ other slopes, while there are only $q - 3$ other points $(i, \sigma(i))$. Hence, only $q - 3$ of the lines $\ell(\mathbf{x}, m)$, $m \in \mathbb{F}_q^* \setminus \{m_j\}$ actually have a point of Γ_σ besides $\mathbf{x}$ on them, leaving one line with no other point. This slope we call $m_j^\perp$.

To see the second claim, we need only apply the previous proposition with $\mathcal{M}_3 = \{m_j\}$, $\mathcal{M}_1 = \{m_j^\perp\}$. $\square$

Problem 3. *Can an edge of* $\mathcal{C}(\Gamma_\sigma)$ *be isolated, i.e., its three points lonely?*

If not, then we may conclude that $\|\Gamma_\sigma\| \geq 2q/5$ by a simple counting argument. Indeed, we have the following.

Proposition 7. *If* $T \subset \Gamma_\sigma$ *of cardinality* t *induces no isolated edges and no isolated points in* $\mathcal{C}(\Gamma_\sigma)$, *then*

$$\|\Gamma_\sigma \cap T\| \geq \frac{2t}{5}.$$

Proof. Clearly, it suffices to prove a lower bound of $2/5$ on the quotient $\|H\|/t$ for all connected hypergraphs H on t vertices with at least 2 edges and so that each edge contains at least three points. Note that, for any such H, it is possible to order its edges $e_1, \ldots, e_k$ so that each e_j, $2 \le j \le k$, intersects e_i for some $i < j$. Define

$$a_j = \left| e_j \setminus \bigcup_{i<j} e_i \right|.$$

Then $a_1 = |e_1|$ and $a_j \le |e_j| - 1$ for $j > 1$. Since $t = \sum_{j=1}^k a_j$,

$$\begin{aligned}
\frac{\|H\|}{t} &= \frac{\sum_{e \in H}(|e|-1)(|e|-2)}{2\sum_{j=1}^k a_j} \\
&\ge \frac{\sum_{e \in H}|e|^2 - 3\sum_{e \in H}|e| + 2|H|}{2(\sum_{e \in H}|e| - |H| + 1)} \\
&\ge \frac{|H|^{-1}(\sum_{e \in H}|e|)^2 - 3\sum_{e \in H}|e| + 2|H|}{2(\sum_{e \in H}|e| - |H| + 1)}
\end{aligned}$$

by Cauchy-Schwarz. Let $\epsilon = \sum_{e \in H} |e|$ and $h = |H|$. Then this reduces to

$$\frac{\|H\|}{t} \ge \frac{\epsilon^2 - 3\epsilon h + 2h^2}{2h(\epsilon - h + 1)}$$

which it is easily checked over $h \ge 2$, $\epsilon \ge 3h$ attains its minimum of $2/5$ at $h = 2$, $e = 6$. $\qquad\square$

Define an *isolated matching* of a hypergraph H to be a set of (disjoint) isolated edges, i.e., a matching which is not intersected by any other edges of H.

Lemma 8. *For every isolated matching $M \subset \mathcal{C}(\Gamma_\sigma)$, there are at least $2|M| + t - \|\Gamma_\sigma\| - 1$ slopes $m \in \mathbb{F}_q^*$ so that there exists an $\ell(\mathbf{x}, m)$ whose intersection with Γ_σ consists of exactly one point of T, where $t = |T|$.*

Proof. Suppose that $M \subset \mathcal{C}(\Gamma_\sigma)$ has exactly $(q - t)/3$ edges, leaving a set $T = \Gamma_\sigma \setminus \bigcup M$ of t other vertices. Let S_1 be the set of slopes of all collinear triples contained in T, so that $|S_1| \le \|T\|$. Note that every edge of $\mathcal{C}(\Gamma_\sigma)$ containing a point of T in fact contains three points of T, since M is an isolated matching. Consider the set of slopes S_2 which occur in M, or whose

8

negative is an element of M. Then $|S_2| \leq 2|M| = 2(q-t)/3$, and $|S_1 \cup S_2| \leq \|T\| + 2(q-t)/3$. Hence, there exist at least

$$|\mathbb{F}_q^* \setminus (S_1 \cup S_2)| \geq q - 1 - \left(\|T\| + \frac{2(q-t)}{3} \right) = \frac{q + 2t}{3} - \|T\| - 1,$$

slopes $m \in \mathbb{F}_q^*$ which are not elements of $S_1 \cup S_2$. Since $\|T\| = \|\Gamma_\sigma\| - (q-t)/3$, this quantity is at least

$$\frac{q + 2t}{3} - \left(\|\Gamma_\sigma\| - \frac{q-t}{3} \right) - 1 = \frac{2q + t}{3} - \|\Gamma_\sigma\| - 1$$
$$= 2|M| + t - \|\Gamma_\sigma\| - 1.$$

If a line ℓ of slope m intersects M in a point $\mathbf{x}$, then, since $m \notin S_2$ and M is an isolated matching, ℓ cannot contain more than one other point of Γ_σ. Furthermore, since $-m \notin S_2$, ℓ contains at least one other point of Γ_σ. Now, if a line ℓ of slope m does not intersect M but *does* intersect T, then it cannot intersect T in three or more points, since $m \notin S_1$. Altogether, this means that every line of slope m intersects Γ_σ either (a) nowhere, (b) at two points of M, (c) at one point of T, (d) at two points of T, or (e) at one point of M and one point of T.

The conclusion follows from the observation that the quantities $|\ell(\mathbf{x}, m) \cap \Gamma_\sigma|$ of types (a),(b),(d), and (e) are even, while q, which is odd, satisfies

$$q = \sum_{\ell \text{ of slope } m} |\ell \cap \Gamma_\sigma|.$$

Hence there is an odd number of lines of type (c). $\qquad\square$

Now, for a point $\mathbf{x} \in \Gamma_\sigma$, define $\|\mathbf{x}\|$ to be the sum over all maximal collinear subsets $\{\mathbf{y}_1, \ldots, \mathbf{y}_k\} \subset \Gamma_\sigma$ of the quantity $\phi(k) = (k-1)(k-2)/(2k)$.

Lemma 9. *Suppose that there are B distinct lines $\ell(\mathbf{x}, m)$, $m \in \mathbb{F}_q^*$, through a point $\mathbf{x} \in \Gamma_\sigma$ so that*

$$\ell(\mathbf{x}, m) \cap \Gamma_\sigma = \mathbf{x}.$$

Then $\|\mathbf{x}\| \geq B/3$.

Proof. Suppose there are A lines $\ell(\mathbf{x}, m)$ so that $|\ell(\mathbf{x}, m) \cap \Gamma_\sigma| \geq 3$, and let $a_1, \ldots, a_A$ denote the quantities $|\ell(\mathbf{x}, m) \cap \Gamma_\sigma|$. Among the slopes $m \in \mathbb{F}_q^*$,

9

there are B lines $\ell(\mathbf{x}, m)$ so that $|\ell(\mathbf{x}, m) \cap \Gamma_\sigma| = 1$, and $q - 1 - (A + B)$ lines so that $|\ell(\mathbf{x}, m) \cap \Gamma_\sigma| = 2$. On the other hand, since $|\Gamma_\sigma| = q$,

$$q - 1 - (A + B) + 2A \leq q - 1 - (A + B) + \sum_{j=1}^{A}(a_j - 1) = |\Gamma_\sigma \setminus \{\mathbf{x}\}| = q - 1,$$

whence $A \leq B$ and $\sum_{j=1}^{A} a_j = 2A + B$. Let $\delta = B/A \geq 1$. Then, by the convexity and monotonicity of ϕ on $[3, \infty)$,

$$\|\mathbf{x}\| = \sum_{i=1}^{A} \phi(a_i) \geq A\phi(\sum_{i=1}^{A} a_i) = A\phi\left(\frac{2A + B}{A}\right) = \frac{B}{\delta}\phi(2 + \delta)$$
$$= \frac{B}{\delta}\frac{\delta(\delta + 1)}{2(\delta + 2)} = \frac{B(\delta + 1)}{2(\delta + 2)} \geq \frac{B}{3}.$$

$\square$

Corollary 10. *Suppose that $M \subset \mathcal{C}(\Gamma_\sigma)$ is an isolated matching, $T = \Gamma_\sigma \setminus \bigcup M$, and there are R distinct lines ℓ intersecting T so that $|\ell \cap \Gamma_\sigma| = 1$. Then*

$$\|T\| \geq R/3.$$

Proof. It is easy to see that, since M is an isolated matching,

$$\|T\| = \sum_{\mathbf{x} \in T} \|\mathbf{x}\|.$$

The conclusion then follows immediately from Lemma 9. $\square$

Theorem 11. $\|\Gamma_\sigma\| \geq \frac{5q-1}{14}$.

Proof. Suppose that $M \subset \mathcal{C}(\Gamma_\sigma)$ is a maximal isolated matching. Define $T = \Gamma_\sigma \setminus \bigcup M$ and $t = |T|$. By Lemma 8, there are at least $2|M| + t - \|\Gamma_\sigma\| - 1$ distinct lines which intersect T in at least one point. Therefore, by Corollary 10, $\|T\| \geq 2|M|/3 + t/3 - \|\Gamma_\sigma\|/3 - 1/3$. Then

$$\|\Gamma_\sigma\| = \|T\| + |M| \geq \frac{5|M|}{3} + \frac{t}{3} - \frac{\|\Gamma_\sigma\|}{3} - \frac{1}{3},$$

whence

$$\frac{q - t}{3} = |M| \leq \frac{4\|\Gamma_\sigma\|}{5} - \frac{t}{5} + \frac{1}{5}$$

10

from which we may deduce

$$t \geq \frac{5q - 3}{2} - 6\|\Gamma_\sigma\|.$$

Since M is maximal, T does not induce any isolated edges. Therefore, by Proposition 7,

$$\|\Gamma_\sigma\| = |M| + \|T\| \geq \frac{q - t}{3} + \frac{2t}{5} = \frac{q}{3} + \frac{t}{15} \geq \frac{q}{2} - \frac{2\|\Gamma_\sigma\|}{5} - \frac{1}{10}.$$

Rearranging, we find that

$$\|\Gamma_\sigma\| \geq \frac{5q - 1}{14}.$$

$\square$

4 Semipermutations

Throughout this section, we assume that σ is a function from $\mathbb{F}_q$ to itself which repeats exactly one value, i.e., $|\sigma(\mathbb{F}_q)| = q - 1$, and so that there are no isolated points in $\mathcal{C}(\Gamma_\sigma)$. Fix the following notation: $\sigma(\mathbb{F}_q) = \mathbb{F}_q \setminus \{b\}$, $\sigma(z_1) = \sigma(z_2) = a$ for some $z_1, z_2 \in \mathbb{F}_q$, $z_1 \neq z_2$.

Proposition 12. *If $\mathbf{x} = (j, \sigma)$ is lonely for $j \neq z_1, z_2$, then*

1. *There is one slope $m_j \in \mathbb{F}_q^*$ so that $\ell(\mathbf{x}, m_j)$ contains three points of Γ_σ, one slope $m_j^\perp$ so that $\ell(\mathbf{x}, m_j^\perp)$ contains one point of Γ_σ, and all other slopes $m \neq m_j, m_j^\perp$ are such that $\ell(\mathbf{x}, m)$ contains two points.*

2. *$m_j^\perp = -m_j \cdot (\sigma(j) - b)/(\sigma(j) - a)$.*

Proof. The proof of (1) is identical to that of Proposition 6. To see the second claim, note that

$$\prod_{i \neq j} \frac{\sigma(j) - \sigma(i)}{j - i} = \prod_{m \in \mathbb{F}_q^*} m \cdot \frac{m_j}{m_j^\perp} = -\frac{m_j}{m_j^\perp}.$$

On the other hand, the numerator of the left-hand side is given by

$$\prod_{i \neq j} (\sigma(j) - \sigma(i)) = \prod_{s \in \mathbb{F}_q^*} s \cdot \frac{\sigma(j) - a}{\sigma(j) - b} = -\frac{\sigma(j) - a}{\sigma(j) - b},$$

11

and the denominator is $\prod_{s \in \mathbb{F}_q^*} s = -1$, whence

$$-\frac{m_j}{m_j^\perp} = \frac{\sigma(j) - a}{\sigma(j) - b},$$

and the conclusion follows. $\qquad\square$

Proposition 13. *If* $\mathbf{x} = (j, a)$ *is lonely for* $j = z_1$ *(or* z_2*), then*

1. *There is one slope* $m_j \in \mathbb{F}_q^*$ *so that* $\ell(\mathbf{x}, m_j)$ *contains three points of* Γ_σ*, two slopes* $m_{j1}^\perp, m_{j2}^\perp$ *so that each* $\ell(\mathbf{x}, m_{ji}^\perp)$ *contains one point of* Γ_σ*, and all other slopes* $m \neq m, m_{j1}^\perp, m_{j2}^\perp$ *are such that* $\ell(\mathbf{x}, m_j)$ *contains two points.*

2. $m_{j1}^\perp m_{j2}^\perp = m_j \cdot (a - b)/(z_2 - z_1)$ *(respectively,* $m_j \cdot (a - b)/(z_1 - z_2)$*).*

Proof. Write $j' = z_2$ (respectively, z_1). Of the $q - 1$ lines $\ell(\mathbf{x}, m)$ through $\mathbf{x}$, $m \in \mathbb{F}_q^*$, there is exactly one m_j for which $|\ell(\mathbf{x}, m_j) \cap \Gamma_\sigma| = 3$ (since $\mathbf{x}$ is lonely), and all other intersections contain 1 or 2 points. Since $q - 2$ of the points of Γ_σ are counted exactly once by these quantities – that is, all of Γ_σ except $\mathbf{x}$ and (j', a) – we have

$$q - 2 = \sum_{m \in \mathbb{F}_q^*} (|\ell(\mathbf{x}, m_j) \cap \Gamma_\sigma| - 1) = 2 + |\{m : |\ell(\mathbf{x}, m_j) \cap \Gamma_\sigma| = 2\}|.$$

This, in turn, implies that there are $q - 4$ slopes $m \in \mathbb{F}_q^*$ so that $\ell(\mathbf{x}, m_j) \cap \Gamma_\sigma$ consists of exactly two points, and the remaining $(q - 1) - (q - 4) - 1 = 2$ directions $m_{j1}^\perp, m_{j2}^\perp$ contain only one point: $\mathbf{x}$.

For the second claim, note that

$$\prod_{i \neq j, j'} \frac{\sigma(j) - \sigma(i)}{j - i} = \prod_{m \in \mathbb{F}_q^*} m \cdot \frac{m_j}{m_{j1}^\perp m_{j2}^\perp} = -\frac{m_j}{m_{j1}^\perp m_{j2}^\perp}.$$

On the other hand, the numerator of the left-hand side is given by

$$\prod_{i \neq j, j'} (\sigma(j) - \sigma(i)) = \prod_{s \in \mathbb{F}_q^*} s \cdot \frac{1}{a - b} = \frac{1}{b - a},$$

and the denominator is $\prod_{s \in \mathbb{F}_q^*} s/(j - j') = 1/(j' - j)$, whence

$$\frac{m_j}{m_{j1}^\perp m_{j2}^\perp} = \frac{j' - j}{a - b},$$

and the conclusion follows. $\qquad\square$

12

Lemma 14. *For every isolated matching $M \subset \mathcal{C}(\Gamma_\sigma)$, there are at least $2|M| + t - \|\Gamma_\sigma\| - 3$ slopes $m \in \mathbb{F}_q^*$ so that there exists an $\ell(\mathbf{x}, m)$ whose intersection with Γ_σ consists of exactly one point of T, where $t = |T|$.*

Proof. Suppose that $M \subset \mathcal{C}(\Gamma_\sigma)$ has exactly $(q - t)/3$ edges, leaving a set $T = \Gamma_\sigma \setminus \bigcup M$ of t other vertices. Let S_1 be the set of slopes of collinear triples in T, so that $|S_1| \leq \|T\|$. Note that every edge of $\mathcal{C}(\Gamma_\sigma)$ containing a point of T in fact contains three points of T, since M is an isolated matching. Consider the set S_2 of slopes m_j and $m_j^\perp$ which occur in M. Then $|S_2| \leq 2|M| = 2(q - t)/3 + 2$ (the 2 arising in case z_1 or $z_2 \in \bigcup M$), and $|S_1 \cup S_2| \leq \|T\| + 2(q - t)/3 + 2$. Hence, there exist at least

$$|\mathbb{F}_q^* \setminus (S_1 \cup S_2)| \geq q - 1 - \left(\|T\| + \frac{2(q - t)}{3} + 2 \right) = \frac{q + 2t}{3} - \|T\| - 3,$$

slopes $m \in \mathbb{F}_q^*$ which are not elements of $S_1 \cup S_2$. Since $\|T\| = \|\Gamma_\sigma\| - (q - t)/3$, this quantity is at least

$$\frac{q + 2t}{3} - \left(\|\Gamma_\sigma\| - \frac{q - t}{3} \right) - 3 = \frac{2q + t}{3} - \|\Gamma_\sigma\| - 3$$

$$= 2|M| + t - \|\Gamma_\sigma\| - 3.$$

If a line ℓ of slope m intersects M in a point $\mathbf{x}$, then, since $m \notin S_2$ and M is an isolated matching, ℓ cannot contain more than one other point of $\mathcal{C}(\Gamma_\sigma)$. Furthermore, since $-m \notin S_2$, ℓ contains at least one other point of Γ_σ. Now, if a line ℓ of slope m does not intersect M but *does* intersect T, then it cannot intersect T in three or more points, since $m \notin S_1$. Altogether, this means that every line of slope m intersects Γ_σ either (a) nowhere, (b) at two points of M, (c) at one point of T, (d) at two points of T, or (e) at one point of M and one point of T.

The conclusion follows from the observation that the quantities $|\ell(\mathbf{x}, m) \cap \Gamma_\sigma|$ of types (a),(b),(d), and (e) are even, while q, which is odd, satisfies

$$q = \sum_{\ell \text{ of slope } m} |\ell \cap \Gamma_\sigma|.$$

Hence there is an odd number of lines of type (c). $\qquad\square$

Lemma 15. *Suppose that there are B distinct lines $\ell(\mathbf{x}, m)$, $m \in \mathbb{F}_q^*$, through a point $\mathbf{x} \in \Gamma_\sigma$ so that*

$$\ell(\mathbf{x}, m) \cap \Gamma_\sigma = \mathbf{x}.$$

If $\mathbf{x} \neq z_1, z_2$, then $\|\mathbf{x}\| \geq B/3$; otherwise $\|\mathbf{x}\| \geq (B - 1)/3$.

13

Proof. If $\mathbf{x} \neq z_1, z_2$, the proof is identical to that of Lemma 9. So, suppose that $\mathbf{x} = z_1$ or z_2, and that there are A lines $\ell(\mathbf{x}, m)$ so that $|\ell(\mathbf{x}, m) \cap \Gamma_\sigma| \geq 3$. Let $a_1, \ldots, a_A$ be the quantities $|\ell(\mathbf{x}, m) \cap \Gamma_\sigma|$. Among the slopes $m \in \mathbb{F}_q^*$, there are B lines $\ell(\mathbf{x}, m)$ so that $|\ell(\mathbf{x}, m) \cap \Gamma_\sigma| = 1$, and $q - 1 - (A + B)$ lines so that $|\ell(\mathbf{x}, m) \cap \Gamma_\sigma| = 2$. On the other hand, since $|\Gamma_\sigma| = q$,

$$q - 1 - (A + B) + 2A \leq q - 1 - (A + B) + \sum_{j=1}^{A}(a_j - 1) = |\Gamma_\sigma \setminus \{z_1, z_2\}| = q - 2,$$

whence $A \leq B - 1$ and $\sum_{j=1}^{A}(a_j - 1) = 2A + B - 1$. Let $\delta = (B - 1)/A \geq 1$. Again, by the convexity and monotonicity of ϕ on $[3, \infty)$,

$$\|\mathbf{x}\| = \sum_{i=1}^{A} \phi(a_i) \geq A\phi(\sum_{i=1}^{A} a_i) = A\phi\left(\frac{2A + B - 1}{A}\right) = \frac{B - 1}{\delta}\phi(2 + \delta)$$
$$= \frac{B - 1}{\delta} \cdot \frac{\delta(\delta + 1)}{2(\delta + 2)} = \frac{(B - 1)(\delta + 1)}{2(\delta + 2)} \geq \frac{B - 1}{3}.$$

$\square$

Corollary 16. *Suppose that $M \subset \mathcal{C}(\Gamma_\sigma)$ is an isolated matching, $T = \Gamma_\sigma \setminus \bigcup M$, and there are R distinct lines ℓ intersecting T so that $|\ell \cap \Gamma_\sigma| = 1$. Then $\|T\| \geq (R - 2)/3$.*

Proof. Since M is an isolated matching,

$$\|T\| = \sum_{\mathbf{x} \in T} \|\mathbf{x}\|.$$

The conclusion then follows immediately from Lemma 15, since at most two of the $\mathbf{x} \in T$ are z_1 or z_2. $\square$

Theorem 17. $\|\Gamma_\sigma\| \geq \frac{5q - 5}{14}$.

Proof. Suppose that $M \subset \mathcal{C}(\Gamma_\sigma)$ is a maximal isolated matching. Define $T = \Gamma_\sigma \setminus \bigcup M$ and $t = |T|$. By Lemma 14, there are at least $2|M| + t - \|\Gamma_\sigma\| - 3$ distinct lines which intersect T in at least one point. Then, by Corollary 16, $\|T\| \geq 2|M|/3 + t/3 - \|\Gamma_\sigma\|/3 - 5/3$, from which it follows that

$$\|\Gamma_\sigma\| = |M| + \|T\| \geq \frac{5|M|}{3} + \frac{t}{3} - \frac{\|\Gamma_\sigma\|}{3} - \frac{5}{3},$$

14

whence
$$\frac{q-t}{3} = |M| \le \frac{4\|\Gamma_\sigma\|}{5} - \frac{t}{5} + 1$$
from which we may deduce
$$t \ge \frac{5q-15}{2} - 6\|\Gamma_\sigma\|.$$
Since M is maximal, T does not induce any isolated edges. Therefore, by Proposition 7,
$$\|\Gamma_\sigma\| = |M| + \|T\| \ge \frac{q-t}{3} + \frac{2t}{5} = \frac{q}{3} + \frac{t}{15} \ge \frac{5q}{12} - \frac{\|\Gamma_\sigma\|}{5} - \frac{1}{4}.$$
Rearranging, we find that
$$\|\Gamma_\sigma\| \ge \frac{5q-5}{14}.$$

<div style="text-align:right">□</div>

5 Questions

First of all, although it is clear that an understanding of the size of the maximum isolated matching in $\mathcal{C}(\Gamma_\sigma)$ would improve our results, we have made little progress in determining what its extremal values are. The arguments in previous sections can be exploited to achieve the following relatively weak bound on t.

Proposition 18. *Let* $\sigma : \mathbb{F}_q \to \mathbb{F}_q$ *be a permutation. The largest isolated matching M in $\mathcal{C}(\Gamma_\sigma)$ has size less than $(q - \sqrt{q/7})/3$.*

Proof. The number of collinear triples in Γ_σ is, as we have seen, at least $(5q-1)/14$. If there were t points in $T = \Gamma_\sigma \setminus \bigcup M$, the number of collinear triples spanned by these points is at most $\binom{t}{2}/3 = t(t-1)/6$. Hence,
$$\frac{q-t}{3} + \frac{t(t-1)}{6} \ge \frac{5q-1}{14}.$$
Solving for t yields
$$t \ge \frac{3}{2} + \sqrt{\frac{q}{7} + \frac{51}{28}} > \sqrt{\frac{q}{7}}.$$

<div style="text-align:right">□</div>

15

As mentioned before, we do not even know if it is possible to have a single isolated edge in $\mathcal{C}(\Gamma_\sigma)$. Of course, the biggest outstanding problem at this time is closing the remaining $q/7$ gap both in the minimum number of collinear triples of a permutation and the size of the smallest Besicovitch set in $\mathbb{F}_q^2$.

6   Acknowledgements

References

[1] J. N. Cooper, J. Solymosi, Collinear points in permutations. Ann. Comb. **9** (2005), no. 2, 169–175.

[2] X. W. C. Faber, On the finite field Kakeya problem in two dimensions, preprint.

[3] N. Katz, I. Łaba, and T. Tao, An improved bound on the Minkowski dimension of Besicovitch sets in R^3, Annals of Math. **152** (2000), 383–446.

[4] T. Wolff, Recent Work Connected with the Kakeya Problem. Prospects in mathematics (Princeton, NJ, 1996), 129-162, Amer. Math. Soc., Providence, RI, 1999.